A Simple Technique to Infer the Missing Gravity Wave Drag in the Middle Atmosphere Using a General Circulation Model: Potential Vorticity Budget

MANUEL PULIDO

*Department of Physics, FACENA, Universidad Nacional del Nordeste, Corrientes, and IMIT,
UMI-IFAECI/CNRS, CONICET, Buenos Aires, Argentina*

ABSTRACT

A simple technique to infer the missing momentum forcing in a general circulation model is developed and evaluated. The response of the large-scale dynamic equations to an external momentum forcing presents a nonlocal response in the zonal and meridional wind. On the other hand, the response to the external momentum forcing in the potential vorticity (PV) is a local growing geostrophic mode, so that there is a direct relationship between the external momentum forcing and the response in PV. In this work, this fact is exploited to diagnose the missing momentum forcing in the extratropics using a general circulation model. The capability of the simple technique to estimate a concentrated gravity wave forcing is evaluated. A dynamical model is evolved with prescribed sources and sinks of PV and then the technique is used to estimate these known momentum sources and sinks. PV is found to give a much better diagnostic of gravity wave drag compared to the more traditional zonal wind differences. The technique is also used in a realistic environment, in which the sources and sinks of PV in Met Office analyses are determined. The estimation of this missing forcing with this simple technique is compared with the estimation given by a more complex data assimilation technique developed by Pulido and Thuburn and, in general, a good agreement is found. The simple gravity wave drag estimation technique can be used in an online data assimilation cycle, using the increments of the analysis, and also offline, using a general circulation model and observations.

1. Introduction

The knowledge of the forcing that gravity waves produce on the general circulation of the atmosphere is currently to a large extent unknown—only a few global observational constraints have been derived using budget studies (Murgatroyd and Singleton 1961; Shine 1989; Alexander and Rosenlof 1996, 2003) and data assimilation (Pulido and Thuburn 2006, 2008). The forcing of gravity waves that reach the stratopause and aloft is assumed to be large scale, so that there is a global response of the mean (or background) flow to that forcing, which includes the well-known Murgatroyd circulation (e.g., Holton 1982). This meridional circulation heats the winter mesosphere and cools the summer mesosphere, inverting the expected temperature gradient from radiative grounds.

General circulation models cannot resolve the entire gravity wave spectrum. Part of the spectrum at the limit of the model resolution is misrepresented and part of the spectrum, the smaller-scale waves, is not represented at all in the general circulation models. The small-scale waves transport upward a significant part of the total wave momentum flux (Alexander et al. 2010) so that the Murgatroyd circulation is not well represented in the models. To alleviate this systematic model error, general circulation models represent the effects of the subgrid physical processes in the resolved circulation using parameterizations. Simple physical principles are used in the parameterizations to give an estimation of the effects of the unresolved processes. Gravity wave drag parameterizations (Hines 1997; Scinocca 2003) launch a spectrum of waves at tropospheric heights and propagate them vertically until a saturation criterion is satisfied. At the saturation height, a nonreversible momentum flux divergence is produced by the waves so that the model dynamics is forced. In theory, the parameterizations should reproduce as closely as possible the missing drag due to these unresolved gravity waves. However,

Corresponding author address: Manuel Pulido, Facultad de Ciencias Exactas, UNNE, Av. Libertad 5400, Corrientes 3400, Argentina.
E-mail: pulido@exa.unne.edu.ar

the actual gravity wave drag due to small-scale waves needed to constrain the parameterizations is largely unknown from observations (Alexander et al. 2010).

The missing gravity wave drag field in a general circulation model is not unique—an increase in the horizontal resolution changes the resolved spectrum of waves and, therefore, the missing gravity wave drag is that produced by a narrower spectrum of waves. This implies that the parameters of a gravity wave parameterization should be changed when the resolution of a model is changed. Furthermore, if the physics or the numerics of the model are changed, the small-scale waves may change their representation so that the missing gravity wave drag is expected to change. In this context, the development of techniques that estimate objectively the physical parameters of gravity wave drag schemes is highly desirable for the calibration of general circulation models.

The detection of systematic errors and attribution to a particular physical process and geographic location in a climate model is entirely nontrivial. Model errors once they are produced evolve with the dynamics of the model so that they are flow dependent and interact nonlinearly with other parts of the dynamical system. One approach that has been followed to attack this problem is the use of general circulation models coupled in a data assimilation system (e.g., Phillips et al. 2004; Martin et al. 2010), so that model forecasts may be compared with observations if the model is started from initial conditions that are observationally constrained, at least from a large-scale perspective. Systematic model errors are more easily identified from short-range forecasts since errors are localized around their sources. Taking this statement to the extreme limit, Klinker and Sardeshmukh (1992) examined the initial tendency errors. In this limit (toward the initial conditions), the model errors are local and are decoupled between them. Using this diagnosis to identify errors in the momentum balance, they detected that vertical diffusion and orographic gravity wave drag from the parameterization were too strong in the European Centre for Medium-Range Weather Forecasts (ECMWF) model. Rodwell and Palmer (2007) further developed the initial tendency approach and applied it to a particular region in which they noticed that some a priori reasonable values of the entrainment parameter gave unbalances in the initial tendency, showing that those entrainment parameter values were unrealistic.

A sophisticated technique to estimate gravity wave drag using data assimilation has been developed by Pulido and Thuburn (2005, 2006). The technique makes use of the adjoint equations to trace back the effects of gravity wave drag on the mean flow to the source of those effects. The effects are the response of the mean flow to the forcing produced by the gravity wave drag. The technique is able to identify the time and location in which gravity wave drag is forcing the mean flow. However, this variational data assimilation technique needs the adjoint model to determine the model sensitivity to the forcing. The development of an adjoint model of a general circulation model is a large, time-demanding task and so most of the models do not have such adjoint models. Thus, simpler techniques that do not use the adjoint model are desirable. This is the motivation for the present work to develop a simple model-independent technique to estimate missing gravity wave drag that may be readily applied to any general circulation model.

The time resolution of the observations is an important parameter to determine the kind of response to a given forcing. Because the effects that we want to measure are global, we assume in this study that global observations in the middle atmosphere are available every 6–48 h. This time scale determines the time at which the response is examined. In this time scale, the response of the mean-flow dynamic equations to the forcing produced by the gravity wave breaking is the classical geostrophic adjustment problem (Blumen 1972).

Because of Coriolis effects, the (mean flow) dynamic equations are coupled between them in the extratropics. As is shown in this work, the response to a localized point forcing in the zonal momentum equation is a nonlocal tripole pattern. On the other hand, the response to a localized forcing in the potential vorticity equation is a local feature that grows with time. Furthermore, the forcing to the mean flow produced by the breaking of gravity waves has two effects: a gravity wave generation (known as secondary waves) and a change in the geostrophic balance. The gravity wave mode is not projected on the potential vorticity (PV) and only the geostrophic forcing mode is projected on PV. Based on these dynamical features, in this work we propose to determine the missing gravity wave drag using a potential vorticity budget.

In summary, the simple technique that we propose uses two main points to determine the error produced by the unresolved motions. First, the use of short-range forecasts started from analyses so that the model errors are not spread and time superposed between them. Second, it makes use of knowing the response of the equations to the momentum errors so that a dynamic variable, potential vorticity, is used in the technique in which the response to the momentum errors is local. The technique works optimally if it is applied in a data assimilation cycle.

The paper is organized as follows. Section 2 derives one simplified set of equations following the classical treatment of the geostrophic adjustment problem

(Blumen 1972) where the effects of the gravity wave forcing in the mean flow are captured. In this conceptual framework, the main concepts are explained in analytical terms. Then, the implementation of the simple technique for gravity wave drag diagnosis in a general circulation model is discussed in section 3. The results of the simple technique are given in section 4. First, we start with theoretical experiments in which the dynamical model is forced with a known forcing and evolve the model with this known forcing, and then the evolution of this model simulation at different times is taken as observation. The potential vorticity budget is applied to these synthetic observations. In these theoretical experiments, the accuracy of the technique is easily evaluated against the known true forcing. In section 4d, a real case is examined. The missing gravity wave drag estimated with the simple technique is compared with the results of a data assimilation technique. The conclusions of the work are drawn in section 5.

2. Conceptual framework

One of the effects of the small-scale motions that are not resolved by a numerical model is that they produce sources and sinks of momentum that are not present in the model. These uncaptured sources and sinks of momentum, in turn, produce a model error that evolves following the dynamics of the system. The response of the large-scale atmospheric circulation to momentum forcing is governed by the classical geostrophic adjustment problem, which has been examined in several previous works (e.g., Dickinson 1969; Blumen 1972). In this section, we examine the equations of the geostrophic adjustment focused on the estimation of the missing momentum forcing in a climate model.

To mimic the effects of the unresolved gravity waves in a general circulation model, the motions are separated in two scales. The large-scale motions are those that are well resolved by the model and therefore have spatial scales larger than λ_M, the horizontal model resolution. The model variables are represented by $\overline{u}, \overline{v}, \overline{\phi}, \overline{T}$, where the overbar means a spatial average over λ_M. The under-/unresolved motions are represented by u', v', w', ϕ', T' and are thought to be dominated by gravity waves.

As is well known (e.g., Andrews et al. 1987), the small-scale motions affect the large-scale flow through their momentum flux divergence terms [e.g., $-\rho_0^{-1}\partial_z(\rho_0\overline{u'w'})$, $-\rho_0^{-1}\partial_z(\rho_0\overline{v'w'})$], which are generically represented as X, Y—that is, the zonal and meridional gravity wave momentum flux divergences, respectively. In what follows, they are called zonal and meridional gravity wave drag. The large-scale momentum equations are then written as

$$\partial_t\overline{u}_T + \overline{u}_T\partial_x\overline{u}_T + \overline{v}_T\partial_y\overline{u}_T - f\overline{v}_T + \partial_x\overline{\phi}_T = X, \quad (1)$$

$$\partial_t\overline{v}_T + \overline{u}_T\partial_x\overline{v}_T + \overline{v}_T\partial_y\overline{v}_T + f\overline{u}_T + \partial_y\overline{\phi}_T = Y, \quad (2)$$

where the subscript T is introduced to denote that this is the nature simulation (the "true" state), which includes the effects of the small-scale motions (i.e., the gravity wave momentum flux divergences) so that there is no model error in (1) and (2). Equations (1) and (2) are expressed in a simple framework using an f-plane assumption. The mathematical analysis in this section is done with a simple set of equations to illustrate the main concepts. The derivation could be generalized to the more general primitive set of equations in spherical geometry. Indeed, in the potential vorticity analysis and in the experiments with the global dynamical model shown in section 4, a spherical geometry is considered so that beta effects are included.

The assumption of the two separated scales between the slowly varying flow and the perturbations is essential for the treatment, but it is indeed a needed assumption for the development of any model parameterization. The effects of the unresolved scales are assumed to be expressed as forcing terms in the model equations. In between these two separated scales, there is a gray zone in which the model is not realistically representing the wave motions. Thus, a good parameterization must represent the effects of these gray zone motions in the larger-scale flow as much as possible even when the theoretical foundations (e.g., Bretherton and Garrett 1968) of the wave–mean flow interactions are not valid for these motions.

By $\overline{u}_H, \overline{v}_H, \overline{\phi}_H, \overline{T}_H$, let us represent the solution to the same large-scale equation system, whose momentum equations are (1) and (2), but without gravity wave drag, $X = Y = 0$. The state variables of this system without forcing are denoted with the H subscript and are called background state variables. Thus, the momentum equations are

$$\partial_t\overline{u}_H + \overline{u}_H\partial_x\overline{u}_H + \overline{v}_H\partial_y\overline{u}_H - f\overline{v}_H + \partial_x\overline{\phi}_H = 0, \quad (3)$$

$$\partial_t\overline{v}_H + \overline{u}_H\partial_x\overline{v}_H + \overline{v}_H\partial_y\overline{v}_H + f\overline{u}_H + \partial_y\overline{\phi}_H = 0. \quad (4)$$

In principle, (3) and (4) could be the set of equations that we have discretized to develop our numerical model in a grid of λ_M resolution; because of this, the forcing terms due to the smaller-scale gravity waves in (1) and (2) cannot be directly determined and are unknown. In general, a parameterization tries to reproduce the effects of the unresolved subgrid dynamics on the larger-scale dynamics. The parameterization is generally represented by a forcing term, say $F_x(u_H, v_H, \ldots)$, $F_y(u_H, v_H, \ldots)$,

which should be introduced in the RHS of (3) and (4). This forcing term tries to mimic the unknown (X, Y) gravity wave drag.

The error of our model (3) and (4)—the large-scale equation system without any parameterization—is the difference between the solution of the true system (1) and (2), which does reproduce the effects of the small-scale waves and the state variables of (3) and (4), $\bar{u}_E = \bar{u}_T - \bar{u}_H$, $\bar{v}_E = \bar{v}_T - \bar{v}_H$. In reality, the discretization, the numerical schemes, etc., also introduce errors. However, we assume that the missing gravity wave drag (X, Y) is the dominant source of errors and the other sources may be considered negligible. The error $\bar{u}_E$ is considered a small perturbation to the background field $\bar{u}_H$. The sum of $\bar{u}_H$ and $\bar{u}_E$ gives the total field $\bar{u}_T$.

In what follows we omit the overline to represent the large-scale variables.

The error of our model due to the unresolved waves is governed by the difference between the two set of equations: (1) and (2) minus (3) and (4). If we start from true atmospheric conditions $[u_H(t_0), v_H(t_0), \phi_H(t_0), T_H(t_0)] = [u_T(t_0), v_T(t_0), \phi_T(t_0), T_T(t_0)]$ then at least for a short period of time the error is expected to be small. Therefore, the model error is governed by the linear equations. In this conceptual picture, the model error (u_E, v_E) represents the linear response to the unknown gravity wave forcing (X, Y).

Thus, assuming the response of the forcing is linear, the resulting equations from (1) and (2) minus (3) and (4) are

$$\partial_t u_E + u_H \partial_x u_E + v_H \partial_y u_E - f v_E + \partial_x \phi_E = X, \quad (5)$$

$$\partial_t v_E + u_H \partial_x v_E + v_H \partial_y v_E + f u_E + \partial_y \phi_E = Y. \quad (6)$$

The resulting equation system (5) and (6) is the classical geostrophic adjustment problem; however, note that in the initial conditions of the forced variables, the error of the model at the initial time $u_E(t_0) = u_T(t_0) - u_H(t_0)$, $v_E(t_0) = v_T(t_0) - v_H(t_0)$ is assumed to be zero, so that instead of being the geostrophic adjustment to an initial unbalance, it is the geostrophic response to a given forcing. The changes in the error field are considered of smaller scale than the background field u_H, v_H so that terms involving the horizontal derivatives of u_H, v_H, such as $u_E \partial_x u_H$, are neglected.

To avoid the advection terms, a Galilean transformation to a system moving with u_H, v_H could be conducted as in Zhu and Holton (1987), such as $\tilde{x} = x - u_H t$, $\tilde{y} = x - v_H t$. However, for now we consider negligible the advection effects on the response and then we discuss the effects of this assumption in sections 4c and 4d.

The resulting system of equations for the model error is

$$\partial_t u_E - f v_E = -\partial_x \phi_E + X, \quad (7)$$

$$\partial_t v_E + f u_E = -\partial_y \phi_E + Y, \quad (8)$$

$$\partial_{tz} \phi_E + N^2 w_E = 0, \quad (9)$$

$$\partial_x u_E + \partial_y v_E + (\partial_z - H^{-1}) w_E = 0, \quad (10)$$

where w_E is the vertical velocity perturbation induced by the forcing (X, Y), N is the buoyancy frequency, and H is the height scale.

The system (7)–(10) is a geostrophic adjustment problem in which the variables u_E, v_E, w_E, ϕ_E respond to the forcing (X, Y). As long as the evolution of u_E, v_E, w_E, ϕ_E is linear, the general solution could be expressed with the Green function formulation (e.g., Zhu and Holton 1987). In the context of the estimation of systematic model error, which is the aim of this work, the absence of X and Y terms are the source of errors in the model (i.e., a momentum deficit at some location and time) and u_E, v_E, w_E, ϕ_E is the evolution of the model error. A perfect technique to estimate the error source is an inverse technique that, given the evolution of the state variables u_E, v_E, w_E, ϕ_E, should be able to estimate the X, Y terms that reproduce the given evolution in the state variables.

Eliminating variables in (7)–(10), the resulting equations for u_E, v_E are

$$(\partial_{tt}\mathcal{H} + f^2\mathcal{H} + \nabla^2)\partial_t u_E = \partial_{yy}X - \partial_{xy}Y + \partial_{tt}\mathcal{H}X + f\partial_t\mathcal{H}Y,$$
$$(11)$$

$$(\partial_{tt}\mathcal{H} + f^2\mathcal{H} + \nabla^2)\partial_t v_E = \partial_{yx}X - \partial_{xx}Y + \partial_{tt}\mathcal{H}Y - f\partial_t\mathcal{H}X,$$
$$(12)$$

where the height operator is $\mathcal{H} = (\partial_z - H^{-1})N^{-2}\partial_z$ and $\nabla^2 = \partial_{xx}^2 + \partial_{yy}^2$.

The formal solution to the forced system for (11) and (12) has been shown in Pulido and Thuburn (2005) and will be not repeated here. The solution is formed by three modes; two are inertio-gravity waves governed by the wave operator in (11). These are secondary waves and have been found in observations (Scavuzzo et al. 1998; Woods and Smith 2010). There is a third mode that is a steady geostrophic growing mode (e.g., Pulido and Thuburn 2005).

The response in u_E and v_E to a meridional and zonal forcing is related to the second derivatives of the forcing in (11) and (12) so that a monopolar forcing center produces a nonlocal response in the geostropic mode.

Although we have found the equations using u, v to show the response in these variables, it is convenient and more illuminating to use as variables absolute vorticity $\zeta = \partial_x v - \partial_y u$ and divergence $\delta = \partial_x u + \partial_y v$ instead of u, v (e.g., Blumen 1972; Zhu and Holton 1987). We use as vertical coordinate the potential temperature θ. The resulting equations of the true state for pseudo-density, $\sigma = \rho \partial z / \partial \theta$, and vorticity are

$$\partial_t \sigma_T + \mathbf{\nabla} \cdot (\sigma_T \mathbf{u}_T) = 0, \tag{13}$$

$$\partial_t \zeta_T + \mathbf{\nabla} \cdot (\zeta_T \mathbf{u}_T) = X_\zeta, \tag{14}$$

where $\mathbf{u}_T = (u_T, v_T)$ and $X_\zeta = \partial_x Y - \partial_y X$. A similar equation system but without forcing holds for σ_H and ζ_H.

The hydrostatic system (13) and (14) is closed with the equations for vertical velocity across isentropes $\dot{\theta}$ and divergence evolution (e.g., Pulido and Thuburn 2005). We assume that radiative forcing is known so that no error is introduced in the vertical velocity equation. In the divergence equation, a nonrotational forcing $X_\delta = \partial_x X + \partial_y Y$ may introduce errors, but this nonrotational forcing only projects on the two inertio-gravity wave modes—say, secondary waves—so that it is not able to change the mean state. This point is discussed in more detail in Pulido and Thuburn (2005).

The model error $\sigma_E = \sigma_T - \sigma_H$ and $\zeta_E = \zeta_T - \zeta_H$ is governed by the linear equations; from (13) and (14), we deduce

$$\partial_t \sigma_E + \mathbf{\nabla} \cdot (\sigma_E \mathbf{u}_H + \sigma_H \mathbf{u}_E) = 0, \tag{15}$$

$$\partial_t \zeta_E + \mathbf{\nabla} \cdot (\zeta_E \mathbf{u}_H + \zeta_H \mathbf{u}_E) = X_\zeta. \tag{16}$$

The interaction terms $(u_E \cdot \nabla)\sigma_H$ and $(u_E \cdot \nabla)\zeta_H$ are assumed to be negligible since the large-scale variables are assumed to change slowly compared to model error variables.

From the homogeneous system equations, which represent the model, we have

$$\delta_H = -\sigma_H^{-1} D_t^H \sigma_H = -\zeta_H^{-1} D_t^H \zeta_H, \tag{17}$$

where $D_t^H = \partial_t + \mathbf{u}_H \cdot \mathbf{\nabla}$.

Using (15) and (17) in (16) leads to

$$D_t^H \zeta_E - \frac{\zeta_E}{\sigma_H} D_t^H \sigma_H - \frac{\zeta_H}{\sigma_H} D_t^H \sigma_E + \frac{\zeta_H \sigma_E}{\sigma_H^2} D_t^H \sigma_H = X_\zeta. \tag{18}$$

Regrouping terms in (18) yields
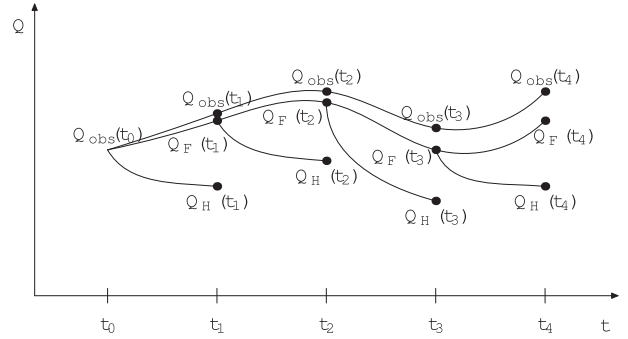
$$\sigma_H D_t^H Q_E = X_\zeta, \tag{19}$$

FIG. 1. Estimation cycle. The observations are represented by $Q_{\mathrm{obs}}(t_i)$; $Q_H(t_i)$ represents the evolution of the model without forcing and $Q_F(t_i)$ is the evolution of the model with the estimated forcing.

where $Q_E = \zeta_E / \sigma_H - \zeta_H \sigma_E / \sigma_H^2$ is the perturbation of potential vorticity (produced by X_ζ).

The sources of error in potential vorticity (i.e., differences in potential vorticity between the nature state and the background state) are associated in (19) to the curl of the forcing for adiabatic motions. Nicely, the potential vorticity in (19) is not governed by the wave operator; only the steady-growing mode remains for this variable. The potential vorticity is also transparent to the divergence of the forcing $X_\delta = \partial_x X + \partial_y Y$. Given a known forcing X_ζ, the evolution of Q_E is readily determined from (19). Note that σ_H and $\mathbf{u}_H$ are assumed to be the "background fields" so that they are known (in practice, they are given by the evolution of the model). The response of Q_E to a X_ζ forcing is entirely trivial and easily invertible since it is local. These properties make the potential vorticity the ideal quantity from which to retrieve the forcing.

3. Implementation

a. Technique details

The implementation of the method to estimate the forcing follows closely a data assimilation cycle (e.g., Pulido and Thuburn 2005). The cycle is illustrated in the diagram shown in Fig. 1. A brief introduction of the proposed technique is given here. Let suppose that we take a period of the cycle of Δt. In the first cycle—say, at time t_0—the initial conditions are taken from reanalysis (or analysis) data. The reanalysis variables are transformed to the model grid. The forecast model without forcing is evolved from t_0 to $t_0 + \Delta t$ [the resulting unforced model state is denoted by $Q_H(t_1)$ in Fig. 1]. Then, the missing forcing is calculated using PV differences between the model forecast and the reanalysis at $t_0 + \Delta t$. The model is evolved again from t_0 to $t_0 + \Delta t$, but now forcing the model with the calculated missing forcing [the resulting forced model state is denoted by

$Q_F(t_1)$ in Fig. 1]. For the second cycle, the model is started from this last forced forecast at $t_1 = t_0 + \Delta t$, $Q_F(t_1)$ and evolved to time $t_1 + \Delta t$ without the forcing term. At $t_1 + \Delta t$, the PV difference between the model forecast and the reanalysis are used to estimate the missing forcing and so on. A more detailed description of the technique is given in what follows of this section.

The numerical model employed in this study is the middle-atmosphere dynamical model of the University of Reading. It represents the full hydrostatic primitive equations in a hexagonal–icosahedral spherical grid (Thuburn 1997). The model covers the middle atmosphere with 16 vertical levels from 100 to 0.1 hPa and the vertical coordinate is potential temperature. A realistic radiative transfer scheme developed by Shine (1987) is used that includes the radiative effects of CO_2, O_3, and H_2O. A gravity wave drag parameterization (Scinocca 2003) is implemented in the model.

The numerical model is started at t_0 taking initial conditions from reanalysis data. Then, the model is evolved from t_0 to $t_1 = t_0 + \Delta t$ without gravity wave drag parameterization. The background state at t_1, $\sigma_H(t_1)$, $u_H(t_1)$, $v_H(t_1)$, $Q_H(t_1)$ is determined in this way. The time-scale Δt in which we evolve the model and determine the forcing using the model and the observations is defined by the available observations that we have to constrain the global-scale problem. These observations are the ones used in global-scale data assimilation systems. A representative time scale for Δt is about 24 h.

The forcing is assumed to be constant in time inside the estimation window Δt considering a time scale of $\Delta t \approx 24$ h. Using (19) and assuming also a constant-in-time σ field,[1] the missing forcing is obtained with

$$X_\zeta = \sigma_{obs}(t_0)[Q_{obs}(t_0 + \Delta t) - Q_H(t_0 + \Delta t)]/\Delta t, \quad (20)$$

where $Q_{obs}(t_0 + \Delta t)$ is the observed PV at $t_0 + \Delta t$ taken from reanalysis data, and $Q_H(t_0 + \Delta t)$ is the PV at $t_0 + \Delta t$ evolved with the numerical model, with the gravity wave drag parameterization switched off, from t_0 to $t_0 + \Delta t$ using as initial conditions the observed state at t_0 (i.e., reanalysis data at t_0). Note that the difference between $Q_{obs}(t_0 + \Delta t)$ and $Q_H(t_0 + \Delta t)$ in (20) includes not only a contribution from the differences between the observed and the background vorticity but also a contribution from

the differences between the observed and the background pseudodensity.

Once the forcing term for the period from t_0 to $t_0 + \Delta t$ is determined with (20), we have two choices to use as initial condition of the model for the next time window $t_0 + \Delta t$, $t_0 + 2\Delta t$: to use the reanalysis data at $t_0 + \Delta t$ or to evolve again the forecast model from t_0 to $t_0 + \Delta t$ but using the determined forcing term, (20), in the model momentum equations. Then, the forced model state at $t_0 + \Delta t$ can be used as initial condition for the next time window. If we use reanalysis data at $t_0 + \Delta t$ as initial condition, then we will be restarting the model with a different state so that the model evolution trajectory will contain jumps at the times $t_0 + n\Delta t$; each of these jumps in the model state is expected to produce some spinup processes. On the other hand, taking the forced model state at $t_0 + n\Delta t$ as initial condition produces a smooth model trajectory as shown in Fig. 1. The evolution of the forecast model with this forcing term, (20), is expected to be close to the observation trajectory, since the forcing is the difference between the observations and the model forecast so that the model momentum equations are nudged to the observed state. A small model drift is expected because the model thermodynamic equation is not nudged to observations.

b. Experimental design

Two sets of experiments were performed. The first set of experiments (cases 1–3) comprises the so-called twin experiments. The second set of experiments is a realistic case in which we use as observations Met Office analyses (case 4). The twin experiments (cases 1–3) are used to evaluate the performance of the technique. In these twin experiments, the model is evolved from a given initial condition $Q_{obs}(t_0)$ with a known forcing X_T, Y_T to generate a set of "synthetic observations," $Q_{obs}(t_i)$ (the forced model state at different times). In the illustration in Fig. 1, the twin experiments have $Q_{obs}(t_i) = Q_F(t_i)$, where $Q_F(t_i)$ is the evolution of the model with the known forcing X_T, Y_T so that there is no model drift in the twin experiments. Using these idealized experiments, we can evaluate the performance of the proposed technique since we know the forcing of the "observed" model state (i.e., X_T, Y_T), which is not known in a real experiment.

For the generation of the synthetic observations with the model, we have to give the initial conditions and the forcing X_T, Y_T. The first experiment, case 1, uses rest-state conditions as initial condition and a zonally symmetric forcing. Case 2 also uses rest-state conditions, but a zonally localized zonal forcing is imposed to the model. The last twin experiment, case 3, uses initial conditions on 1 May 2002 taken from Met Office

---

[1] Equation (19) does support time variations in σ_H. During the course of this study, the assumption of a steady σ field taken in (20) was relaxed using higher-order approximations. However, the evaluation of this higher-order approximation did not lead to any significant improvement in the estimation (not shown).
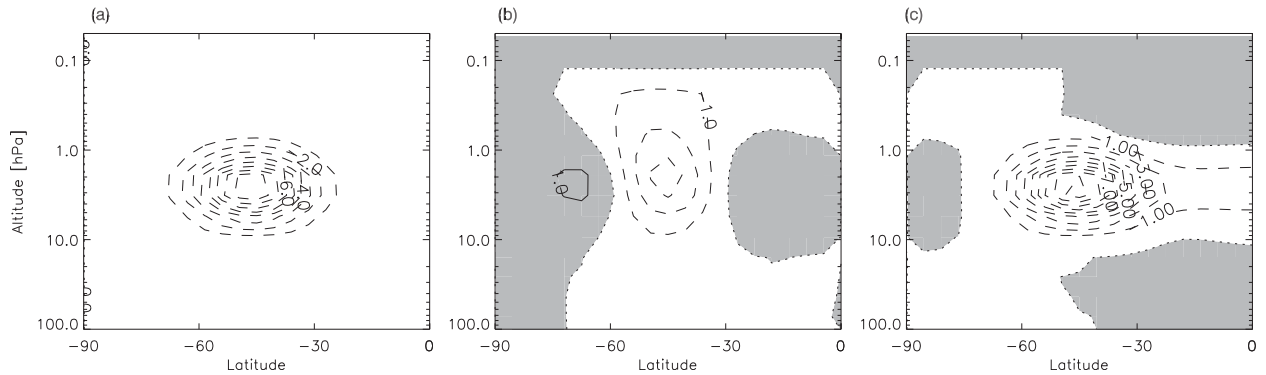
FIG. 2. Experiment with a zonally symmetric forcing in the zonal momentum equation. (a) Prescribed zonal forcing. (b) Zonal forcing estimated using the U estimation. (c) Zonal forcing estimated using the PV estimation. Contour interval is $1 \, \mathrm{m \, s^{-1} \, day^{-1}}$.

analysis. Then the model is evolved, for case 3, with a prescribed localized forcing to generate the synthetic observations in this case.

The application of the forcing estimation technique to the twin experiment cases is exactly the same as the application to a realistic case as explained in section 3a and shown in Fig. 1, except that instead of using reanalysis as observations, we use the synthetic observations.

For an easier interpretation of the results, the estimated X_ζ, from (20), is transformed to obtain the zonal and meridional gravity wave drag components (X, Y) Furthermore, a standard way to evaluate the zonal wind forcing term might be written as

$$X = [u_{\mathrm{obs}}(t_0 + \Delta t) - u_H(t_0 + \Delta t)]/\Delta t. \quad (21)$$

We refer to this as U estimation and use it as a comparison to the PV estimation of the missing forcing.

Finally, a real case, case 4, is examined in section 4d, in which Met Office analyses are used as observations and we estimate the missing gravity wave drag of the model. In this real case, the true missing forcing X_T, Y_T is not known, so we apply the Pulido and Thuburn (2005) data assimilation technique to estimate the forcing and compare with the estimation results given by the technique introduced in this work. The data assimilation approach uses a four-dimensional variational data assimilation technique that has been developed for the same University of Reading model used in this work. The assimilation technique includes the adjoint model and a conjugate gradient minimization module [see Pulido and Thuburn (2005) for further details]. The estimation of gravity wave drag with the data assimilation technique is referred to as DA estimation.

If the gravity wave drag estimation is performed in an operational data assimilation cycle, then the increments contain the information about the difference between observations and the model forecasts. Therefore,

the increments can be used to estimate the missing gravity wave drag. Since the models in the data assimilation systems have implemented a gravity wave parameterization in general, a linear response has to be assumed such that the total missing drag is the sum of the estimated gravity wave drag and the one given by the parameterization. This assumption is evaluated in section 4d.

4. Results

a. Case 1: Estimation of a zonally symmetric forcing: Rest atmosphere

In the case 1 experiment, a zonally symmetric Gaussian forcing located at 1 hPa and at midlatitudes (Fig. 2a) is imposed in the momentum equations of the model. The model is started from rest conditions and from an isothermal atmosphere. The model is evolved adiabatically in this idealized simulation. Therefore, the rest conditions are a solution so that the total field is only the response to the forcing. The simulation is conducted for 24 h. At 24 h, the fields, zonal wind, and PV are taken to perform the diagnosis. Figure 2b shows zonal forcing from U estimation and Fig. 2c shows zonal forcing from PV estimation. A much closer agreement to the true forcing is obtained with the PV estimation. The root-mean-square error (RMSE) of the gravity wave drag using PV estimation is $0.39 \, \mathrm{m \, s^{-1} \, day^{-1}}$ while the RMSE using U estimation is $0.82 \, \mathrm{m \, s^{-1} \, day^{-1}}$ so that the error in the PV estimation is about half the U estimation. This shows that PV field, on which the PV estimation is based, has more dynamical information on the adjustment process than the U field.

The response in the zonal wind at 24 h is shown in Fig. 3a. Three jets are found in the response. Thus, a monopolar zonal forcing has a response in the zonal wind that is a tripolar pattern. This nonlocal behavior is related to (11) in which we found that the u response is
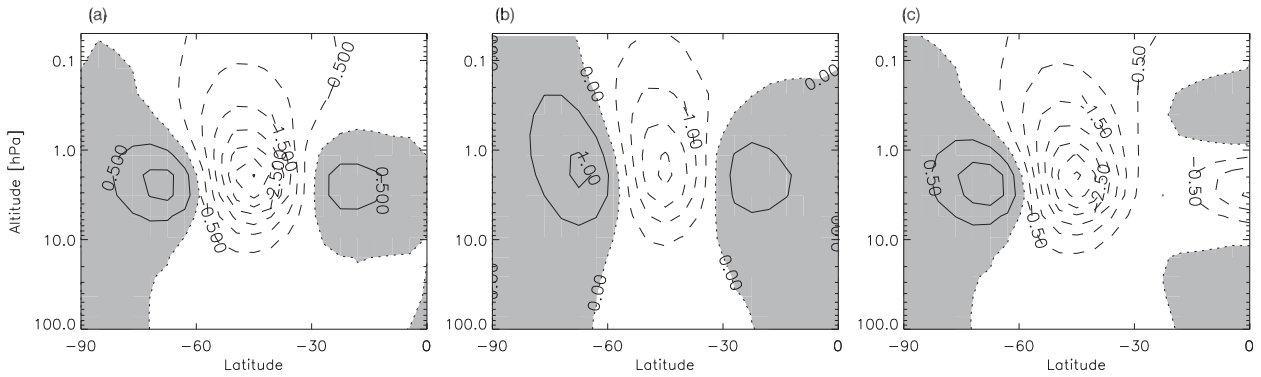
FIG. 3. (a) Response of the zonal flow to the prescribed zonal forcing. (b) Response to the zonal forcing estimated using the U estimation. (c) Response to the zonal forcing estimated using the PV estimation. Contour interval is $0.5\,\mathrm{m\,s^{-1}\,day^{-1}}$.

related to the second meridional derivative of the zonal forcing. The response to the forcing estimated with PV estimation (Fig. 3c) gives a closer agreement with the true response than the forcing from U estimation (Fig. 3b) except close to the equator where geostrophic balance does not hold. In the tropics, the response to a given localized forcing is not governed by geostrophic adjustment so that potential vorticity may not be a good diagnostic variable; indeed, the dynamical equations are decoupled close to the equator.

This experiment shows that the intuitive first-order evaluation of gravity wave drag using zonal wind differences may be imprecise. The response to a localized forcing is governed by an elliptical equation in the zonal flow so that the response is nonlocal in time scales of a few hours. Only for a period of less than 1 h can the response in the zonal flow to the zonal forcing be considered local.

A qualitatively similar tripolar pattern to Fig. 3a in the zonal wind difference was found by McLandress et al. (2012) in the Southern Hemisphere using the analysis increments in the Canadian Middle Atmosphere Model (CMAM). Thus, the missing forcing in the Southern Hemisphere of CMAM may be linked to a concentrated deceleration center. The amplitude of the missing forcing is expected to be stronger than the estimation of the forcing given by U estimation from (21) used by McLandress et al. (2012).

b. Case 2: Estimation of a localized zonal forcing: Rest atmosphere

The second experiment evaluates the evolution of the model forced by a zonally as well as meridionally and vertically localized zonal forcing. This forcing can be divided in two parts: the nonrotational part and the nondivergent part. In what follows, we use a more direct terminology (although less precise), calling the nondivergent part of the forcing "rotational" and the nonrotational part "divergent." Figure 4a shows the localized zonal forcing. Figures 4b and 4c, which show the zonal forcing contributed from the rotational and the divergent part of forcing, are obtained by transforming the zonal prescribed forcing to the rotational and divergence parts of the forcing. Then, the rotational part
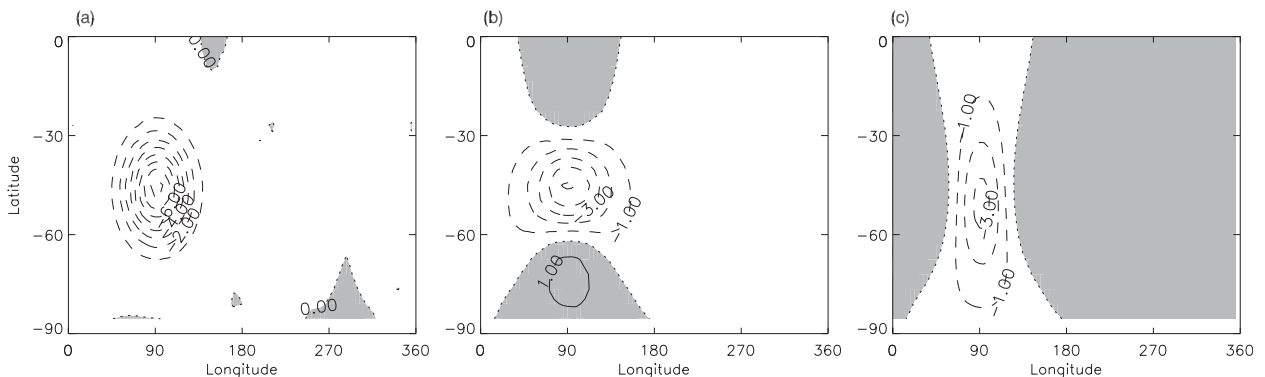

FIG. 4. (a) The horizontally localized prescribed zonal forcing. (b) The zonal forcing contributed from the rotational part of the forcing. (c) The zonal forcing contributed from the divergent part of forcing. Contour interval is $1\,\mathrm{m\,s^{-1}\,day^{-1}}$.
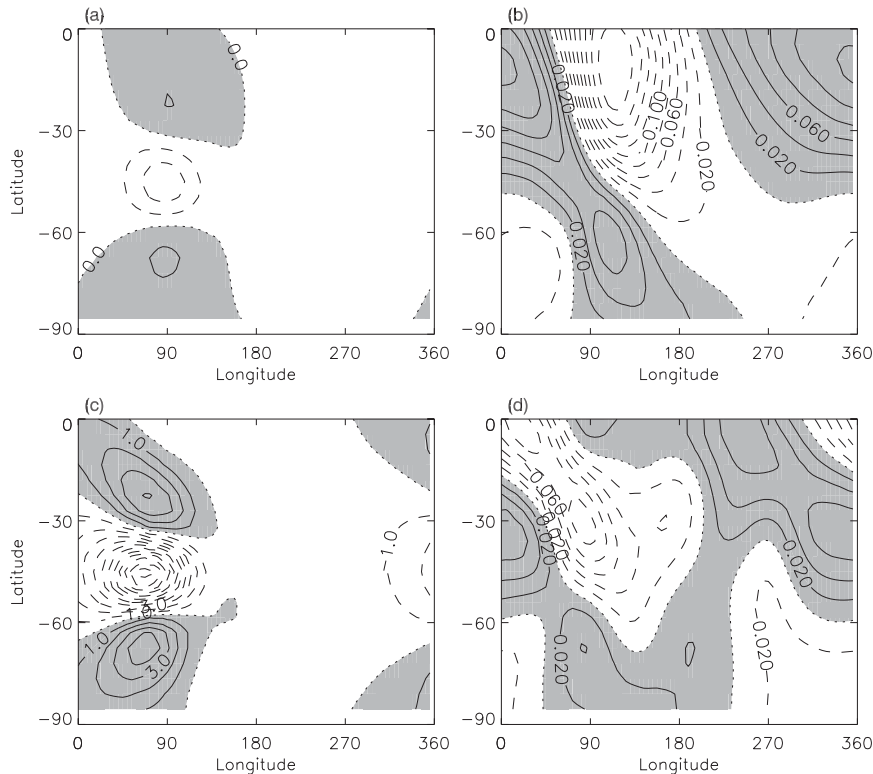
FIG. 5. (a) The zonal wind response at $t = 1$ day to the rotational part of the horizontally localized prescribed zonal forcing (contour interval $1 \, \mathrm{m \, s^{-1}}$). (b) The response at $t = 1$ day to the divergent part of the forcing (contour interval $0.02 \, \mathrm{m \, s^{-1}}$). (c) The response at $t = 5$ days to the rotational part of the forcing [contour interval as in (a)]. (d) The response at $t = 5$ days to the divergent part of the forcing [contour interval as in (b)].

is transformed back to the zonal (and meridional) component of the forcing (Fig. 4b shows the zonal forcing component), and the divergent part is separately transformed back to the zonal (and meridional) component of the forcing (Fig. 4c). Note that the sum of the field in Fig. 4b and the field in Fig. 4c gives the zonal forcing shown in Fig. 4a.

To evaluate the response to the divergent and rotational parts of the forcing, two simulations were conducted: one in which the model is forced by the rotational part of the drag and the other one forced by the divergent part. Figure 5 shows the zonal wind response in the two simulations at $t = 1$ and $t = 5$ days. As seen in section 2, the response to the divergent part of the forcing shown in Figs. 5b and 5d is expected to generate gravity waves and a steady-state nongrowing anomaly (Pulido and Thuburn 2005). On the other hand, the response to the rotational part of the forcing is a growing anomaly (Figs. 5a and 5c), as found in (19). The total response is clearly dominated by the response to the rotational part of the forcing (note that contours intervals are 1 and $0.02 \, \mathrm{m \, s^{-1}}$ for the rotational and the divergent parts, respectively).

The zonal forcing from the curl of the gravity wave drag estimated using PV estimation is shown in Fig. 6a and using U estimation in Fig. 6b. The estimated gravity wave drag shown in Fig. 6 should be compared to the zonal forcing of the true curl of the gravity wave drag shown in Fig. 4b. The maximum error in the estimation of the rotational part of the drag using PV (U) estimation is lower than $0.7 \, (3) \, \mathrm{m \, s^{-1} \, day^{-1}}$. The RMSE in the gravity wave drag estimation is $0.11 \, \mathrm{m \, s^{-1} \, day^{-1}}$ for the PV estimation and $0.18 \, \mathrm{m \, s^{-1} \, day^{-1}}$ for the U estimation.

The response to the divergent part of the forcing is very small and nonlocal so that it is not possible to be estimated precisely using a budget technique. On the other hand, a data assimilation technique is able to estimate the divergent part of the drag in twin experiments (Pulido and Thuburn 2005). Since the divergent part of the forcing cannot change the geostrophic balance and, therefore, it does not affect the general circulation, its estimation does not appear to be as important as the rotational part. Furthermore, the divergent part of the flow is poorly constrained from observations, as is argued in Pulido and Thuburn (2006).
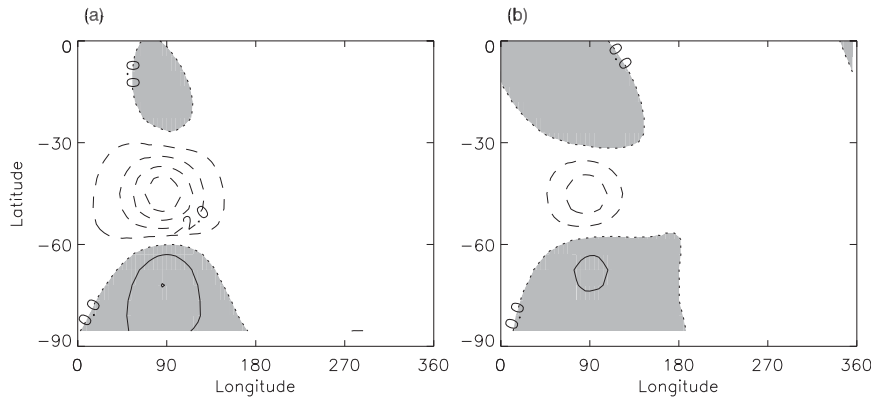
FIG. 6. (a) The estimated zonal forcing contributed from the rotational part of the forcing using PV estimation for the horizontally localized forcing case. (b) The estimated forcing using U estimation. Contour interval is $1 \, \mathrm{m \, s^{-1} \, day^{-1}}$.

c. Case 3: Estimation of a localized nondivergent forcing: Realistic-state atmosphere

The estimation of a localized nondivergent forcing is examined. The response to this prescribed localized forcing is expected from (19) to be local so that it is particularly suitable for evaluation purposes. To evaluate the technique in a more realistic environment than the two previous experiments in which the initial conditions were a rest state, the model is started with initial conditions taken from Met Office analysis for the day of 1 May 2002. The model is then evolved from those initial conditions and with the prescribed localized forcing to obtain the synthetic observed state on 2 May 2002.

Figures 7a and 7d show the zonal wind at 1 hPa and at 90° longitude on 2 May 2002, which was evolved from Met Office analysis on 1 May 2002. The curl of gravity wave drag estimated with the PV estimation is shown in
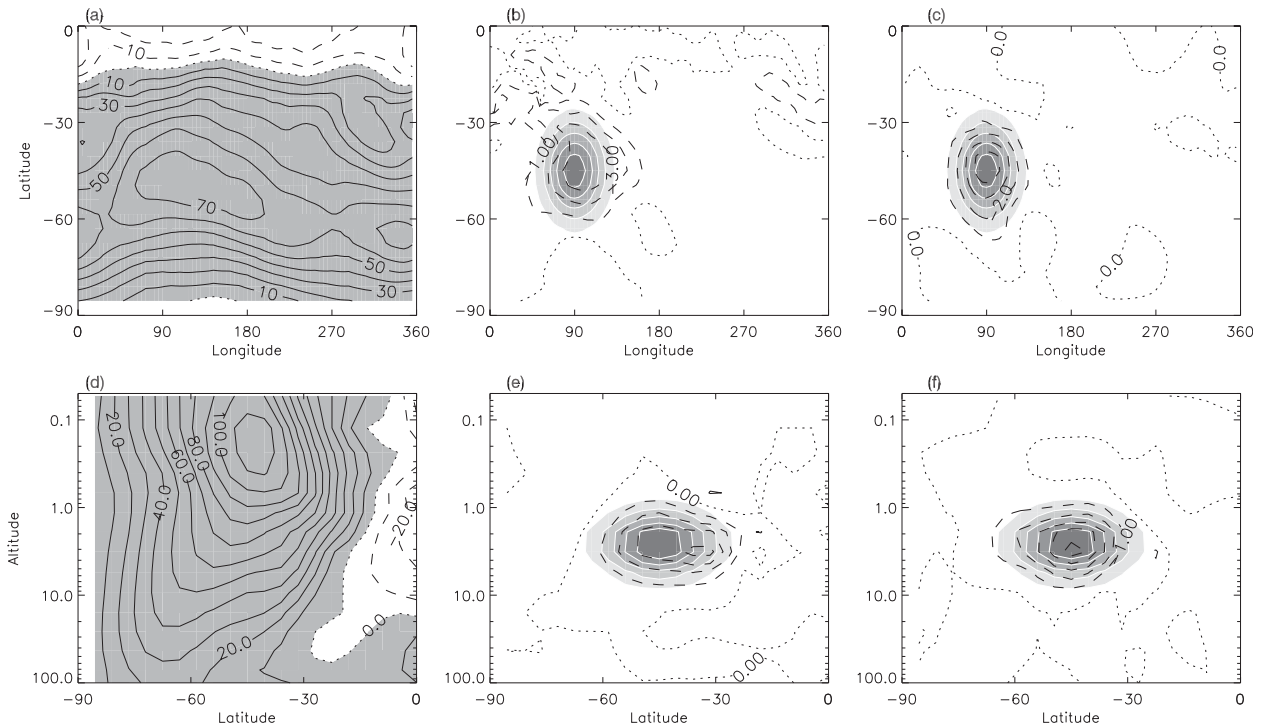

FIG. 7. (a) Cross section at 1 hPa of the zonal wind $(\mathrm{m \, s^{-1}})$ on 2 May 2002 (contour interval $10 \, \mathrm{m \, s^{-1}}$). Prescribed curl of the drag (shading interval $10^6 \, \mathrm{s^{-1}}$) and estimated curl of the drag (contours) using (b) PV estimation and (c) DA estimation. (d)–(f) As in (a)–(c), but a y–z cross section at 90°E.
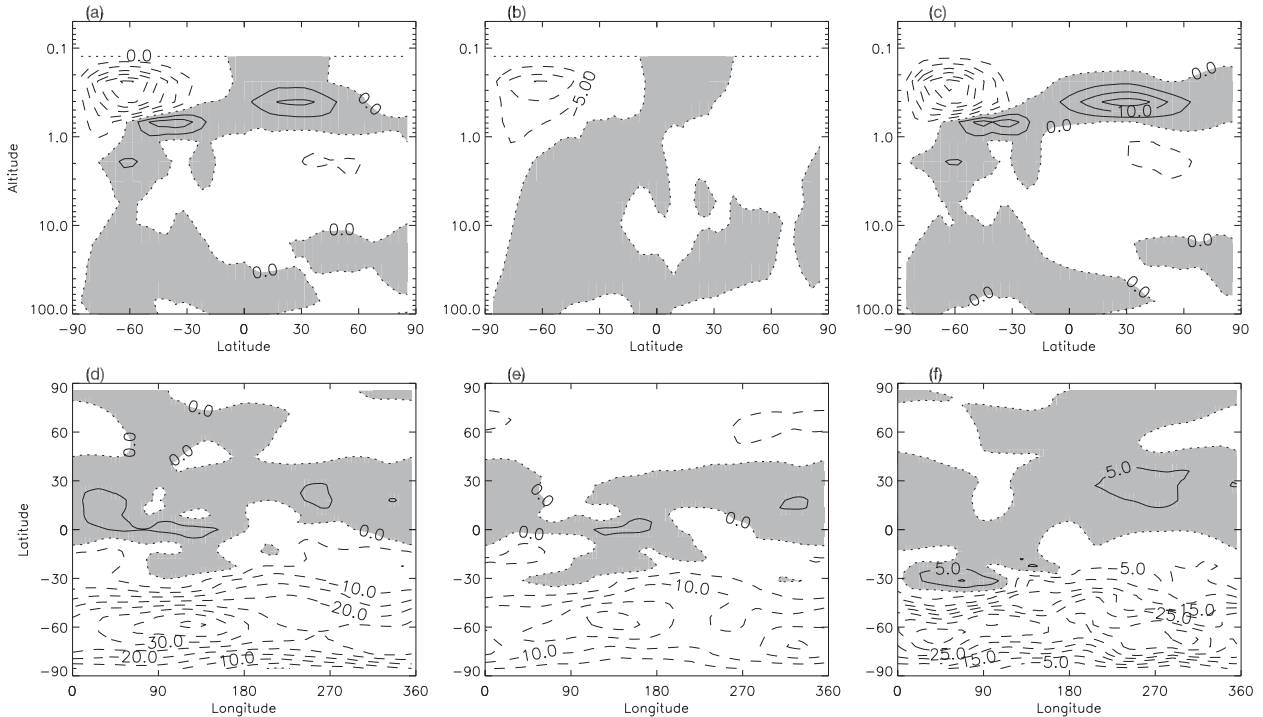
FIG. 8. Weekly mean gravity wave drag estimated (a),(d) using PV estimation, (b),(e) using U estimation, and (c),(f) using data assimilation. (top) Zonally averaged gravity wave drag and (bottom) gravity wave drag at 0.24 hPa. Contour interval is $5\,\mathrm{m\,s^{-1}\,day^{-1}}$.

Figs. 7b and 7e (contours), and the true curl of the gravity wave drag is also shown in the figures (shading). The estimated curl of gravity wave drag is horizontally spread over a slightly larger region than the true curl of gravity wave drag particularly toward the equator and to the east. The estimated values peak at $4.6 \times 10^6\,\mathrm{s^{-1}}$, while the true peak is $6.0 \times 10^6\,\mathrm{s^{-1}}$. A closer agreement is seen in Figs. 7c and 7f between the true curl of gravity wave drag and the one from the DA estimation. The estimated peak in that case is $5.5 \times 10^6\,\mathrm{s^{-1}}$. The contours in the horizontal pattern follow closely to the true gravity wave drag. This experiment shows that the DA estimation is able to trace back the location of the momentum sources even in the presence of strong zonal winds. On the other hand, the PV estimation is affected by advection, but the effect of advection in these strong winds does not appear to be so large for 24-h time windows.

Note that under the approximation used to estimate the rotational part of the forcing, (20), we neglect the linearized advection of forced PV. This is a strong argument to use short time windows. The window we use is 24 h; if shorter window forecasts—say, 6 h—are used, then the drag estimation could improve since the effects are more local. For shorter-range forecasts, the interactions

between the mean flow and the response to the forcing and the advection by the mean flow are smaller.

d. Case 4: Estimation in a real case

In this section, we focus on a real estimation case in the week from 1 to 7 May 2002. Met Office analyses are used as observations. The model is started with initial conditions from the Met Office analysis at 1200 UTC 1 May 2002. The model simulation after 24 h is compared with the Met Office analysis at 1200 UTC 2 May 2002; the curl of the drag during this 24-h time window is estimated from the PV estimation (20). For the second day, the model is started from the previous model simulation, which has been forced by the estimated gravity wave drag. The second gravity wave drag is estimated as the difference between this forecast and Met Office analysis at 1200 UTC 3 May 2002, and so on. This estimation cycle is further discussed in section 3 (see Fig. 1). The estimated weekly averaged zonal gravity wave drag is shown in Figs. 8a and 8d. To evaluate the validity of the PV estimation technique, the U estimation is shown in Figs. 8b and 8e and an estimation of the gravity wave drag using the data assimilation technique as in Pulido and Thuburn (2008) was also conducted during the same week from 1 to 7 May 2002. Figures 8c and 8f
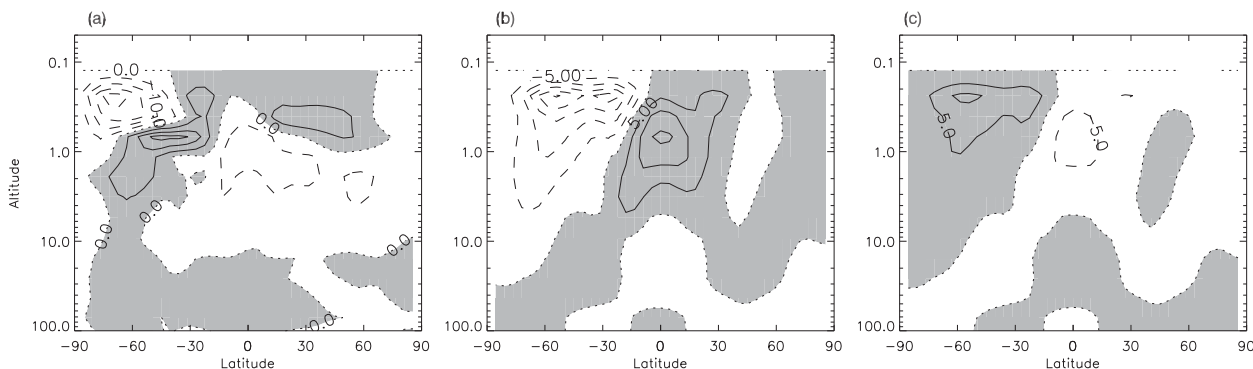
FIG. 9. (a) Weekly mean missing gravity wave drag estimated using PV estimation in an experiment in which the gravity wave drag parameterization is switched on. (b) Weekly gravity wave drag given by the parameterization. (c) Difference in total gravity wave drag between the estimation experiments with and without gravity wave drag parameterization. Contour interval is $5 \, \mathrm{m \, s^{-1} \, day^{-1}}$.

show the weekly averaged zonal gravity wave drag from the data assimilation technique.

We note that we found a better agreement between the gravity wave drag from PV estimation with the gravity wave drag from DA estimation when the forced (nudged) forecasts were used in subsequent windows. If Met Office analyses are used in the subsequent windows as initial conditions, the PV estimation has larger differences compared to the DA estimation, particularly around the equator (not shown). One possible explanation for this is that differences that are not completely corrected in one time window in a smooth model evolution may be corrected in subsequent windows. On the other hand, when the model is restarted every time, the estimation does not have memory from previous estimations and also spinup processes are generated.

The agreement between the PV estimation (Fig. 8, left) and the gravity wave drag from the data assimilation technique (Fig. 8, right) is very good; the location and width of the deceleration center are well detected. The peak magnitude at the center is a little weaker: 30 versus $32 \, \mathrm{m \, s^{-1} \, day^{-1}}$. The acceleration center below the deceleration center—a behavior that is expected from a Lindzen's (1981) type of analysis—is very well captured both in magnitude and shape. In the Northern Hemisphere, the positive deceleration center, centered at 30°N and 0.4 hPa, is also rather well estimated but it is also weaker compared to the data assimilation technique. The peak is 14 versus $18 \, \mathrm{m \, s^{-1} \, day^{-1}}$. This is in agreement with the slight errors in the gravity wave drag estimation that we found in the twin experiment shown in section 4c. The gravity wave drag horizontal patterns shown in Fig. 8d have the maximum deceleration feature, at 60°S, 110°E, located to the west of the maximum feature found in the DA estimation (at 65°S, 40°E; see Fig. 8f). This difference is expected because of advection of the errors by the winter jet (Figs. 7a and 7d) so

that the error sources are expected to be to the east and then the errors are transported by the winter jet (an effect that is not captured by the PV estimation). The PV estimation technique neglects the linear advection (by the winter jet) of forced PV; on the other hand, the four-dimensional data assimilation technique does account for linear advection of forced PV through the adjoint model.

We also conducted an experiment in which the forecasts are produced with the gravity wave drag parameterization (Scinocca 2003) switched on, and the PV estimation is conducted over those forecasts. In this case, the estimated missing gravity wave drag is the one that is not captured by the parameterization (Fig. 9a). The gravity wave drag given by the parameterization is shown in Fig. 9b. The estimated missing gravity wave drag plus the parameterized gravity wave drag gives a similar pattern as in Fig. 8a and d so that the linear superposition assumption is valid in this context. Figure 9c shows the difference between the missing gravity wave drag from Fig. 8a and the sum of estimated missing gravity wave drag and the parameterized gravity wave drag. This experiment shows that standard short-range forecasts that are performed for model evaluation are useful to determine the gravity wave drag that is missing in the parameterization (i.e., to measure the misrepresentation of the gravity wave drag).

5. Conclusions

The use of short-range forecasts that are initially started from analyses constrained by observations, for time lengths from about 6 to 24 h, may be used in the middle atmosphere to determine the momentum sources and sinks of a general circulation model. The effects of these momentum sources are linear and relatively local on these time scales. The momentum sources can

be estimated very precisely using a PV estimation method in the extratropics. It permits us to estimate the non-divergent component of the forcing, which is the component that drives the mean flow. This nondivergent component of the forcing is the only component that is projected on the geostrophic mode. On the other hand, the nonrotational component of the forcing is projected on the inertio-gravity wave modes and is not able to drive the mean flow (e.g., Pulido and Thuburn 2005).

The technique has been extensively examined using twin experiments and has also been applied in a realistic environment initialized with analysis data. In both cases, the estimation of the nondivergent component of the gravity wave drag is excellent compared to the known forcing. In an estimation that uses analysis data as observations, the simple technique has been compared with the gravity wave drag estimation using a data assimilation technique (Pulido and Thuburn 2005), which gives very similar results to the data assimilation estimation, except for the horizontal location of the forcing, which is affected by advection of the response to the forcing.

The evaluation of the gravity wave drag parameterization in the middle atmosphere appears to be easier than similar evaluations in the troposphere where the effects of the parameterizations are coupled. For instance, a precipitation error may be traced back to an error of the convection scheme, the planetary boundary layer parameterization, or the land surface model. On the other hand, in the middle atmosphere the parameters of the radiative scheme are well known so that the main source of model error is due to small-scale motions that cannot be fully captured by the model resolution. These effects of small-scale motions are represented by the nonorographic and orographic gravity wave drag parameterizations. To distinguish between orographic and nonorographic contributions to the estimated missing gravity wave drag does not appear to be trivial apart from the geographical location where the missing gravity wave drag is found, which could be associated with mountains. The other aspect that might contaminate the gravity wave drag estimation is if analyses are not well constrained by observations; in this case, the estimation technique will infer the gravity wave drag given by the parameterization in the data assimilation model. The fact that data assimilation systems [e.g., CMAM (McLandress et al. 2012) and ECMWF (N. Zagar 2013, personal communication)] give important analysis increments (i.e., differences between analysis and model forecasts) where gravity wave drag is expected and estimated by the technique shows that the fields are well constrained by observations there. Using the observed large-scale PV field to determine the forcing requires

both the large-scale wind field (to determine vorticity perturbations induced by the forcing) and temperature field (to determine pseudodensity perturbations induced by the forcing). On the other hand, the U estimation uses only the large-scale observed wind field. Although the PV estimation requires more observational information, the fact that it uses the temperature field to determine the forcing is a positive aspect of the technique since this is the field best constrained by observations.

The technique was also evaluated for the complete month of June 2002 with a time window of 24 h, in which the differences between the nudged forecasts (which are forced by the estimated gravity wave drag) and the observations remain relatively small. For an estimation longer than 1 month, it may be convenient to restart the model from the observations (analysis or reanalysis) every 30 days so that any small drift that is produced after 30 days of estimations and is not well captured by the momentum corrections is eliminated. This small drift between observations $Q_{obs}(t_i)$ and the forced model evolution $Q_F(t_i)$ is illustrated in Fig. 1. Then, we suggest for $i = 30$ to restart the model again with $Q_{obs}(t_{30})$ as the initial condition instead of using $Q_F(t_{30})$; in this way, we eliminate the model drift that is produced in the previous 30 time windows.

We envision two ways of implementing the missing drag diagnosis: in an online data assimilation cycle and in an offline estimation. If the general circulation model is part of a data assimilation system, the increments are available during each data assimilation cycle, which are the differences between the model forecasts (started from the last analysis) and the available observations. Using the increments, the missing nondivergent component of the gravity wave drag for each cycle can be estimated from (20). In this work, we have applied an offline estimation in which the analyses from an external data assimilation system are used as observations; the model is evolved two times in each time window. The first time is a free-run model evolution [denoted by $Q_H(t_i)$ in Fig. 1]. The second model forecast is nudged to the external analysis using the estimations of gravity wave drag to force the model toward the analysis [denoted by $Q_F(t_i)$ in Fig. 1].

Because of its simplicity, the technique can be readily applied to any general circulation model, as opposed to data assimilation techniques, which may require the adjoint model of the general circulation model in the case of a variational data assimilation technique. The use of ensemble Kalman filter techniques to estimate model error is also very promising since they are model-independent techniques (Ruiz et al. 2013). However, to our knowledge, they have not been evaluated to

estimate momentum forcing terms. A related problem—the estimation of the spatial distribution of carbon flux sources—has been successfully obtained with the ensemble Kalman filter (Kang et al. 2012).

The simple technique can be easily applied to the multimodel simulations that are being conducted for the evaluation of the parameterizations of climate models in which short-range forecasts from observed initial conditions are produced—for example, Transpose Atmospheric Model Intercomparison Project (Williams et al. 2013). The data requirements for the gravity wave drag estimation are the short-range forecast state variables (u, v, T, ϕ) and the momentum forcing terms such as the gravity wave drag from the orographic and nonorographic parameterizations (with the frequency of the estimation periods). In the case that the model has the parameterizations switched on during the forecasts, the total missing gravity wave drag is the sum of the parameterization forcing and the estimated missing forcing. The averaged estimated forcing in the middle atmosphere may be interpreted as a lack of tuning the parameters of the gravity wave drag parameterization.

REFERENCES

Alexander, M. J., and K. H. Rosenlof, 1996: Nonstationary gravity wave forcing of the stratospheric zonal mean wind. *J. Geophys. Res.,* **101** (D18), 23 465–23 474.

——, and ——, 2003: Gravity-wave forcing in the stratosphere: Observational constraints from the upper atmosphere research satellite and implications for parameterization in global models. *J. Geophys. Res.,* **108,** 4597, doi:10.1029/2003JD003373.

——, and Coauthors, 2010: Recent developments in gravity wave effects in climate models, and the global distribution of gravity wave momentum flux from observations and models. *Quart. J. Roy. Meteor. Soc.,* **136,** 1103–1124.

Andrews, D. G., J. R. Holton, and C. B. Leovy, 1987: *Middle Atmosphere Dynamics.* Academic Press, 489 pp.

Blumen, W., 1972: Geostrophic adjustment. *Rev. Geophys.,* **10,** 485–528.

Bretherton, F., and C. Garrett, 1968: Wavetrains in inhomogeneous moving media. *Proc. Roy. Soc. London,* **A302,** 529–554.

Dickinson, R. E., 1969: Propagators of the atmospheric motions 1. Excitation by point impulses. *Rev. Geophys.,* **7,** 483–514.

Hines, C. O., 1997: Doppler spread parametrization of gravity-wave momentum deposition in the middle atmosphere. Part 1: Basic formulation. *J. Atmos. Sol. Terr. Phys.,* **59,** 371–386.

Holton, J., 1982: The role of gravity wave induced drag and diffusion in the momentum budget of the mesosphere. *J. Atmos. Sci.,* **39,** 791–799.

Kang, J., E. Kalnay, T. Miyoshi, J. Liu, and I. Fung, 2012: Estimation of surface carbon fluxes with an advanced data assimilation methodology. *J. Geophys. Res.,* **117,** D24101, doi:10.1029/2012JD018259.

Klinker, E., and P. Sardeshmukh, 1992: The diagnosis of mechanical dissipation in the atmosphere from large-scale balance requirements. *J. Atmos. Sci.,* **49,** 608–627.

Lindzen, R. S., 1981: Turbulence and stress owing to gravity wave and tidal breakdown. *J. Geophys. Res.,* **86** (C10), 9707–9714.

Martin, G. M., S. F. Milton, C. A. Senior, M. E. Brooks, S. Ineson, T. Reichler, and J. Kim, 2010: Analysis and reduction of systematic errors through a seamless approach to modeling weather and climate. *J. Climate,* **23,** 5933–5957.

McLandress, C., T. G. Shepherd, S. Polavarapu, and S. R. Beagley, 2012: Is missing orographic gravity wave drag near 60°S the cause of the stratospheric zonal wind biases in chemistry–climate models? *J. Atmos. Sci.,* **69,** 802–818.

Murgatroyd, R. J., and F. Singleton, 1961: Possible meridional circulations in the stratosphere and mesosphere. *Quart. J. Roy. Meteor. Soc.,* **87,** 125–135.

Phillips, T. J., and Coauthors, 2004: Evaluating parameterizations in general circulation models: Climate simulation meets weather prediction. *Bull. Amer. Meteor. Soc.,* **85,** 1903–1913.

Pulido, M., and J. Thuburn, 2005: Gravity wave drag estimation from global analyses using variational data assimilation principles. I: Theory and implementation. *Quart. J. Roy. Meteor. Soc.,* **131,** 1821–1840.

——, and ——, 2006: Gravity wave drag estimation from global analyses using variational data assimilation principles. II: A case study. *Quart. J. Roy. Meteor. Soc.,* **132,** 1527–1543, doi:10.1256/qj.05.43.

——, and ——, 2008: The seasonal cycle of gravity wave drag in the middle atmosphere. *J. Climate,* **21,** 4664–4679.

Rodwell, M. J., and T. N. Palmer, 2007: Using numerical weather prediction to assess climate models. *Quart. J. Roy. Meteor. Soc.,* **133,** 129–146.

Ruiz, J., M. Pulido, and T. Miyoshi, 2013: Estimating parameters with ensemble-based data assimilation: A review. *J. Meteor. Soc. Japan,* **91,** 79–99.

Scavuzzo, C., M. Lamfri, H. Teitelbaum, and F. Lott, 1998: A study of the low-frequency inertio-gravity waves observed during the Pyrénées Experiment. *J. Geophys. Res.,* **103** (D2), 1747–1758.

Scinocca, J., 2003: An accurate spectral nonorographic gravity wave drag parametrization for general circulation models. *J. Atmos. Sci.,* **60,** 667–682.

Shine, K., 1987: The middle atmosphere in the absence of dynamical heat fluxes. *Quart. J. Roy. Meteor. Soc.,* **113,** 603–633.

——, 1989: Sources and sinks of zonal momentum in the middle atmosphere diagnosed using the diabatic circulation. *Quart. J. Roy. Meteor. Soc.,* **115,** 265–292.

Thuburn, J., 1997: A PV-based shallow-water model on a hexagonal–icosahedral grid. *Mon. Wea. Rev.,* **125,** 2328–2347.

Williams, K. D., and Coauthors, 2013: The Transpose-AMIP II experiment and its application to the understanding of southern ocean cloud biases in climate models. *J. Climate,* **26,** 3258–3274.

Woods, B. K., and R. B. Smith, 2010: Energy flux and wavelet diagnostics of secondary mountain waves. *J. Atmos. Sci.,* **67,** 3721–3738.

Zhu, X., and J. R. Holton, 1987: Mean fields induced by local gravity-wave forcing in the middle atmosphere. *J. Atmos. Sci.,* **44,** 620–630.